Interurban Wage and Rent differences: the Value of Air Quality and Crime in Chile.[*]

by

Felipe Vásquez Lavín

Department of Agricultural and Resource Economics, University of California, Berkeley, 311 Giannini Hall.

---

[*] I am grateful to Maximilian Auffhammer, Sofia Villas-Boas, Rui Huang, James Manley and Susan Stratton for their useful comments on the draft version of this paper. I would like to blame them for any errors which remain but my good manners compel me to absolve them of any guilt and take full responsibility myself.

Abstract

In this article I examine the existence of compensating differentials in the housing and labor markets using data from a national survey in Chile. Contrasting with earlier literature I consider the simultaneous character of decisions made by individual in both the labor and housing market. Emphasis is placed in the joint estimation of both wage and rent equations taking into account the correlation between the error terms in these two equations. A first round of estimations includes an OLS, a SURE and a simultaneous Tobit model. I tested the endogeneity of some amenities of interest using a Hausman test. The hypothesis of endogeneity can not be rejected. Therefore the estimations are corrected for endogeneity using 2SLS and 3SLS. Finally I computed welfare measures for two of the most important problems nowadays in Chile, crime and air pollution. Results show that i) increments in crime rate or air pollution have a negative impact in welfare and ii) welfare measures estimates and their variance are highly biased when the simultaneity is not considered.

1 Introduction

As an extension of the theory of compensating wage differentials, hedonic wage theory postulates that when workers have the same productivity and the market is perfectly competitive, differences in wages will reflect differences in the non-wage characteristics of the jobs. In particular, different levels of amenities, ceteris paribus, should be reflected in the gap of equilibrium wages obtained by otherwise "equal jobs and workers" (Rosen 1979, 1986).

These amenities include characteristics that are completely exogenous to the economic agents such as climate components (precipitation, temperature, humidity, etc), as well as other characteristics that are a result of the agents' decisions such as urban conditions and environmental quality. When an individual chooses a place to live or work, he or she indirectly affects other aspects of the community such as population density, average education and income, number of cars in the location, etc. In the same way, when firms make location decisions, they also affect some characteristics of the neighborhood through emissions to the environment, labor demand, etc.

In particular I anticipate that conditions like air quality and crime rates generate a wage differential among different locations, to the extent that these characteristics vary among regions. Thus in places where environmental conditions are considered to be better one would expect that, all other factors equal, wages should be permanently lower than in those regions where conditions are considered to be worse. This is the essence of the theory of hedonic wages.

This idea can be traced back to hedonic price theory, which looks at the differences in housing prices to obtain a welfare measure associated with an environmental amenity. In this literature, the value of an environmental amenity can be approximated through the estimation of a cross section price equation, which can be a function of location specific amenities (say air pollution and crime), house and other neighborhood characteristics. Studies that specifically address the effect of air quality on housing prices are Ridker and Henning (1967), Anderson and Crocker (1971), Freeman (1971, 1974) and Smith and Deyak (1975); crime has been studied by Roback (1982) and Deller et al. (2001) among others.

In equilibrium the marginal implicit price for the amenity will equate the consumers' marginal willingness to pay, and hence the derivative of the housing price function with respect to the environmental amenity can be interpreted as its marginal (implicit) price (Rosen, 1974; Freeman, 1979; Palmquist, 1991).

One difference between hedonic wage and price theory is that while for property value models the derivative of the price function with respect to the environmental amenity can be interpreted as the marginal price for this amenity, this is no longer so in the wage model. The reason is that from the worker's perspective the housing and job decisions are connected. Generally, a decision to move to a job in another city will be at the same time a decision to change residence to that city. But, if an environmental amenity does have an effect on the job and housing location characteristics, then these decisions will not be independent. For example if the prospect of living and working in a clean environment increases the desirability of a job in a certain city, probably the demand for housing services will be high in this city and therefore housing rents will also be high. Thus the worker should consider the trade-off between having a "better" job and paying a more expensive rent. This is why the job and housing decisions are modeled together in the formal hedonic wage models (Rosen, 1979; Marin and Psacharopoulos, 1982; Roback, 1988; Hoehn et al. 1987). The principal result of this procedure is that the marginal willingness to pay for a higher environmental quality will depend on both the willingness to accept a lower wage and the willingness to spend more on housing (Freeman, 2003).

From an estimation perspective one should model at least a two-equation system with a wage and a housing price equation. Since Roback (1982) developed the model for two simultaneous equation systems, many authors have applied it to estimate the value of some environmental amenities. Nevertheless, most of them have totally ignored the possible correlation between the error terms in these equations. The usual practice has been to estimate two separate equations, one equation for wage and another for rent, using OLS or some nonlinear Box-Cox transformation. Some examples of this kind of applications include papers by Roback (1982, 1989), Hoehn et al. (1987, 1988), and Deller and Ottem (2001) among others. Clark and Kahn

(1988) calculate a willingness to pay function for an environmental amenity through the estimation of a hedonic wage model using only an equation for earnings. However, they do not consider the simultaneous decision problem of work and housing location, that is to say, they estimate only a partial willingness to pay function.

Joint estimation of the two equation system is important not only because of the possibility of a more efficient estimation of the standard errors of the parameters, but also for the reason that a joint estimation allows us to obtain the covariance between estimators in both equations. This covariance is required when we desire to evaluate the significance of the welfare measure associated with an amenity. As we will see, the welfare measure can be expressed as a linear combination of parameters from both the wage and rent equations; accordingly its variance could be easily calculated provided that we have the covariance between parameters. Some authors simply do not provide any estimation of the welfare measure's variance, and others ignore the covariance effect (Hoehn et al. 1988 and Blomquist et al. 1988), Finally, Deller and Ottem (2001) estimate the variance of the marginal value using a Monte Carlo simulation, based on the asymptotic properties of the OLS estimators.

Another relevant issue in hedonic prices is the selection of the functional form for the hedonic equation. Most of the literature in this regard is concerned with the house hedonic equation. Among the most common functional forms used for the house hedonic equation are the linear, the quadratic, the log-log, the semi-log and the Box-Cox transformation. Some simulation evidence shows that a simple Box-Cox transformation provides a good approximation for the price equation (Bender, 1982; Cropper, Deck and McConnell, 1988; Palmquist and Danielson, 1989; Mc Connell, 1995). Other general discussions include papers by Graves et al. (1988) and Palmquist et al. (1997) among others. However, the issue of functional form in the joint model of rent and wages has not been fully discussed in the literature. Moreover, Cassel and Mendelsohn (1985) shed some light on the cost of using a Box-Cox transformation when the researcher is interested in estimating the price for characteristics or predicting the changes in prices. In general the Box-Cox transformation will reduce the accuracy of coefficients and produce bad prediction results.

This article has a threefold purpose. First, this paper addresses the problem of simultaneity in the system of equations. As I will discuss, this problem can be thought of as a seemingly unrelated model with a Tobit structure in one of the equations (the wage equation). Following this idea, I use three different estimation models: 1) Heckman's two step model, where in the second step I use OLS for each equation separately; 2) Heckman's two step model, where in the second step I use a joint estimation of the wage and rent equations (SURE model); and 3) a Tobit simultaneous equation system. The last two models enable me to estimate the covariance among parameters in both equations. These three methods of estimation are extended to consider the possible endogeneity of some explanatory variables.

The second goal is to apply the hedonic wage theory to the Chilean case. As far as I am aware, this is the first attempt to apply this model to the Chilean case. Figueroa and Lever (1992) estimated a hedonic price model using information from the house market in Santiago. Even though the model was used successfully to predict the value of houses, it did not include any environmental variables. Thus, the marginal willingness to pay for amenities was not estimated. In the same way, another paper by Wunder and Gutierrez (1992) proposed a strategy for evaluating and estimating Box-Cox models. Once again environmental variables were not considered. Finally, Figueroa, Rogat and Firinguetti (1996) estimated only an hedonic price function and obtained welfare measures on environmental amenities (for a 50% reduction in air pollution).

So far, the lack of reliable environmental information and labor market data in Chile have made it difficult to estimate a hedonic wage model as suggested in this paper. In that sense this application tests the quality of the available information on environmental amenities and tells us something about how much we can really learn from the data. In addition, that this application allows us to obtain welfare measures with direct implications for public environmental policies. This process is particularly relevant today given that the two most significant problems nowadays in Chile are crime and air pollution.

The last objective is the estimation of welfare measures for both air quality and crime in the Chilean market, including the calculation of their standard errors using the parameters' variance and covariance obtained in the models estimation.

The next section discusses the basic theoretical underpinnings of the paper, i.e. the theory of hedonic wages and how they can be related to the environment. Then I present the econometrics issues related with the model and discuss means of analyzing the empirical evidence. A discussion about the data used in this study follows. Finally, I present the econometric results and some conclusions close the paper.

2 Methodological approach

I draw the theoretical framework directly from that presented in Roback (1982). Using this setup I assume that workers consume a composite good X, at a price that will be a numeraire, and that they rent l^c units of land at a price r. The workers' income is made up of wages, w, and non-wage income, I. Thus, workers in this theoretical setup try to solve the following mathematical problem:

$$Max \quad U(X, l^c; s), \tag{1}$$

$$s.t. \quad w + I = X \ + l^c * r,$$

where s stands for amenities. Equilibrium on the worker's (consumer) side, can be characterized by:

$$V(w, r : s) = c, \tag{2}$$

where $V(w, r; s)$ is the indirect utility function associated with this problem and c is a constant. I set $V_s > 0$, implying that utility increases as consumption of the amenity increases, and $V_r < 0$, $V_w > 0$ as usual.

On the other hand, firms maximize profits given a constant returns to scale production function with land (l^p) and total number of workers (N) as inputs. In equilibrium the unit cost function for firms follows

$$C(w, r; s) = 1, \tag{3}$$

with $C_r = l^p/X > 0$, $C_w = N/X > 0$ and $C_s > 0$ if the amenity is unproductive and $C_s < 0$ for a productive amenity.

Totally differentiating equations (2)and (3) we can solve for the comparative static multipliers of wage and rent with respect to an amenity s.

$$\frac{\partial w}{\partial s} = \frac{1}{|A|}\left(-V_s C_r + V_r C_s\right) < 0, \tag{4}$$

$$\frac{\partial r}{\partial s} = \frac{1}{|A|}\left(-V_w C_s + V_s C_w\right) \lesseqgtr 0, \tag{5}$$

where $|A| = V_w C_r - V_r C_w > 0$. Given that the amenity is unproductive ($C_s > 0$), an increase in the level of the amenity will unambiguously decrease the equilibrium wage level. On the other hand, the effect on housing rents will be ambiguous depending on the relative effects of firms' and workers' decisions (Roback, 1982). Workers prefer locations with higher levels of amenities, so they increase the supply of labor in locations with higher s as they increase the demand for land. On the other hand firms prefer locations with lower s (because s is unproductive), decreasing both labor and land demand in places with higher s. In the case of wages the effects go in the same direction: lower salaries are observed in places with higher s, but the effect on rents is ambiguous, depending on the relative size of the demand and supply effects.

Figure 1 shows two possible scenarios after a change in the level of an amenity. Both the indirect utility function and the unit cost function are shown. For the initial level of s the equilibrium is given by the combination (r, w). After a change in the unproductive amenity, the rent and wage must go up in order to keep the same level of utility for the worker. Moreover, the cost for the firm must go down in order to satisfy the optimality condition of the firm. Thus, the final results could be at $(w', r') \ll (w, r)$, or at (w'', r''), with $w'' < w$ and, $r'' > r$.

Equations (2) and (3) cannot be directly observed, but we can observe (4) and (5). Roback (1982) shows that from these equations and using Roy's identity we can deduce the following formula for estimating the implicit price (p_s^*) associated with the amenity s:

$$p_s^* = -\frac{V_s}{V_w} = l^c \frac{\partial r}{\partial s} - \frac{\partial w}{\partial s}. \tag{6}$$

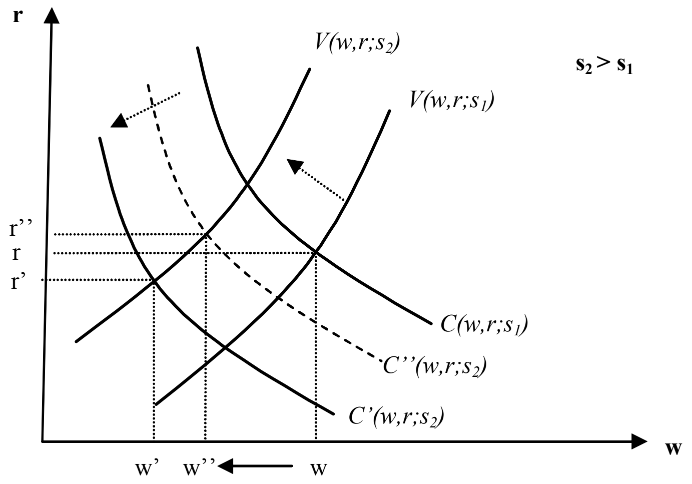

Figure 1

This model has been commonly used in the literature. For example, Berger et al. (2003) applied this model to estimate quality of life indices using information on emerging labor and housing markets in Russia. Deller and Ottem (2001) estimate an index of life quality for Wisconsin counties, based on the marginal value of crime rates. A more sophisticated model has been proposed by Hoehn et al. (1987) and Blomquist et al (1988), where they incorporate other elements such as intraurban commuting cost, total amount of land available and population density into the model. In their model the effects on the endogenous variables depend on a complex interaction of factors. In particular the effects on $\frac{\partial w}{\partial s}$ and $\frac{\partial r}{\partial s}$ are ambiguous and depend on several assumptions about the effect of changes in the amenity level on firms and individuals, along with agglomeration effects. Ultimately, though, the estimation of the welfare measure is essentially the same as Roback (1982).

3 Econometric estimation

In order to obtain $\frac{\partial w}{\partial s}$ and $\frac{\partial r}{\partial s}$ I estimate the outcome of both the labor and housing markets using the data described in next section. The model includes two equations.

$$y_{1i} = \alpha x_i + \varepsilon_{1i}, \tag{7}$$

$$y_{2i}^{*} = \beta z_i + \varepsilon_{2i}, \tag{8}$$

y_{2i}^{*} is a latent variable linked to wages by the following formulation

$$y_{2i} = \begin{cases} y_{2i}^{*} = y_{2i} \;\; if \quad y_{2i}^{*} > 0 \\ \quad 0 \;\; otherwise. \end{cases} \tag{9}$$

The first equation models the rents where $i \in \{1, ..., n\}$ indexes people included in the sample, $\alpha = (\alpha_1, ..., \alpha_k)$ is a vector of k parameters, x_i is the corresponding vector of explanatory variables for observation i, and ε_{1i} are the stochastic disturbances. The second equation models the wage where $\beta = (\beta_1, ..., \beta_l)$ is a vector of l parameters, z_i is the corresponding vector of explanatory variables for observation i and ε_{2i} are stochastic disturbances. This model is a seemingly unrelated simultaneous equations system where the first equation is linear and the second is a Tobit model. If we assume that the disturbances are *i.i.d.* with

$$\begin{bmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \right), \tag{10}$$

then following Maddala (1983) the likelihood function can be expressed as

$$L = \prod_{y_2>0} f(y_1, y_2) \prod_{y_2=0} \int_{-\infty}^{0} f(y_1, y_2^*) dy_2^*, \tag{11}$$

where $f(y_1, y_2)$ is the joint density function of y_1 and y_2 when $y_2 > 0$. This model was fully developed by Nelson and Olson (1978) in the context of a simultaneous-equation model with limited dependent variables. Unlike Nelson and Olson we do not have to worry about the reduced form of the model since no endogenous variables are included in the right hand side of system (7) and (8). Using a change of variables rule the likelihood function can be written as:

$$L = \prod_{y_2>0} f(y_{1i} - \alpha x_i, y_{2i} - \beta z_i) \prod_{y_2=0} \int_{-\infty}^{0} f(y_{1i} - \alpha x_i, y - \beta z_i) dy. \tag{12}$$

This model can be easily estimated with conventional econometrics packages like LIMDEP. In appendix 2, I describe some details of the likelihood function.

Notice that this model implicitly assumes that the mechanism determining whether or not a person is employed is the same as the mechanism that determines the wages. In order to avoid this constraint we could modify the model. First define a new variable

$$y_{3i}^* = \gamma w_i + \varepsilon_{3i}, \tag{13}$$

related to the wages, y_{2i}, and employment status,y_{3i}, by

$$\begin{aligned} y_{3i}^* &\leq 0 \Rightarrow y_{3i} = 0 \text{ and } y_{2i} = 0 \\ y_{3i}^* &> 0 \Rightarrow y_{3i} = 1 \text{ and } y_{2i} = y_{2i}^*. \end{aligned}$$

That is, the mechanism determining employment differs from the mechanism determining wages, and $\gamma = (\gamma_1, ..., \gamma_m)$ is a vector of m parameters and w_i is the corresponding vector of explanatory variables for observation i. Finally, ε_{3i} are the stochastic disturbances.

We can assume that the disturbances are identically and independently distributed as

$$\begin{bmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \\ \varepsilon_{3i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix} \right).$$

Clearly the likelihood function associated with this model is more complicated that the first one. Fortunately this model is similar to a type 3 Tobit model discussed in Amemiya (1985). Therefore we can use Heckman's two-step estimator (Heckman, 1976). That is, we estimate a participation equation (equation 13) using a probit model and including the whole sample. This equation permits us to obtain the inverse Mill's ratio (λ_i)[1] , which is used as another explanatory variable in the wage equation. In the second step GLS is applied simultaneously to both equations but

---

[1] See appendix 2 for details on Heckman's two-step model.

only for individuals with positive earnings. The equation for the second step would be

$$\begin{aligned} y_{1i} &= \alpha x_i + \varepsilon_{1i}, & (14) \\ y_{2i} &= \beta z_i + \beta_\lambda \lambda_i + \varepsilon_{1i}. & (15) \end{aligned}$$

As Amemiya (1983) pointed out, more efficient estimators are obtained using GLS for simultaneous equations.

In my literature review, I could not find any application of the wage hedonic model that takes into account both the selection bias associated with the earning equations and the intrinsic simultaneity of the model. In a slightly different setting, Berger et al. (2003) used Heckman's two-step model to correct the rent equations but not the wage equation, arguing that as many households did not report any rent, OLS will provide biased estimators. Despite the fact that this could also be true with my data I will not follow this approach now.

The last issue considered in this econometric section involves the estimation of welfare measures and their variance. Given the available data, my estimated equations differ from those by Roback because I estimate the logarithm of the hourly wage instead of the total wage income, and the logarithm of per capita land rental cost instead of the per area unit rental. Therefore I have to adjust equation (6), and the new welfare measure is given by

$$p_s^* = a\frac{\partial \ln a}{\partial s} - L_w\frac{\partial \ln w}{\partial s}. \tag{16}$$

In this formula a is the per capita rental cost and L_w is wage income for the same individual. Note that $\frac{\partial \ln a}{\partial s}$ and $\frac{\partial \ln w}{\partial s}$ can be obtained directly from the estimations as the estimators of the corresponding parameters.

At least two different approaches can be used for estimating the implicit marginal price of the amenity. One of them is to estimate this value for the representative worker; that is, to evaluate the expression in (16) at the sample means of the variables. As I assume these sample means to be fixed, I get the implicit price of the amenities as a linear combination of two random variables: the estimators of the parameters of the rental costs and wages. This will be the estimate of the **marginal price of the amenities** faced by the representative individual. Alternatively, I can estimate the implicit price for each individual and take the average. This will be the **average of the marginal price of the amenities**. In this model, the formula in (16) results in a linear expression; thus, both approaches lead to similar results.

The variance of the welfare measure can be expressed with the following formula:

$$\sigma_p^2 = \begin{bmatrix} b_a & b_w \end{bmatrix} \begin{bmatrix} \sigma_{aa} & \sigma_{aw} \\ \sigma_{wa} & \sigma_{ww} \end{bmatrix} \begin{bmatrix} b_a \\ b_w \end{bmatrix} \tag{17}$$

where $b_i,\ i \in \{a, w\}$ are the estimates of the coefficients of rental costs and wages, respectively. The matrix in the middle of this equation is the matrix of variances and covariance of these estimators. Traditionally researchers have ignored the estimation of the variance of the welfare measure or they have implicitly assumed that $\sigma_{wa} = 0$. In the estimation section I will compare the result of this assumption with the results obtained assuming $\sigma_{wa} \neq 0$.

4 The data

I use several sources of data. First, individuals' characteristics were mainly obtained from the CASEN (National Survey on Socioeconomic Characterization) survey for the year 1998. This survey includes information on a representative sample of households throughout Chile, including 188,000 individuals. It includes data on family structure, housing, health, education, labor and income conditions of the families as well as individual data on wages, self-reported housing rents, labor and residential characteristics, and some firm characteristics.

I define several dummy variables which allow me to control for a household's head (1 for head), marital status (1 for married), and sex (1 for male), among others. A number of variables are related to the characteristics of the dwellings, including a dummy for existence of adequate sanitary installations (WC), and a system of dummy variables describing the quality of walls, floor, and roof (better quality has a 1 indicator).

Several variables describe the job of the individual being interviewed. There is a dummy that reflects enrollment in Social Security (SS). Additionally, I discriminate between tenured, permanent jobs and temporary positions, contract (with and without), and working for large and micro enterprises compared to a baseline medium size enterprise. I use dummies to specify nine possible employment sectors, with "industry" as the baseline. Wages are measured on an hourly basis and rents have been reduced to a per capita basis, dividing by the number of members in the household.

In this study I use data on only 30 municipalities out of a total of 225 in the whole dataset. This is necessary because these 30 are the only municipalities with available environmental data.

Environmental and amenity data were taken from other sources. Air quality is defined in terms of particulate material (PM10). These data were collected both from SESMA, the Environmental Health Service for Metropolitan Regions, and from the CONAMA, the National Committee for the Environment. Monitoring stations run by these organizations produced these data.

Finally, data on other amenities and location characteristics comes from the SINIM, the National Municipal Information System. These variables include per capita government expenditures, poverty rate, potable water coverage, population density, cleaning services expenditures, and others. A dummy variable identifies municipalities in the Chilean Capital (SAN). The number of reported crimes per thousand inhabitants in each municipality comes from the INE, National Institute of Statistics. Finally, I also used the INE as the source for data on unemployment rates and numbers of vehicles in municipalities. Table 1 in Appendix 1 presents the main characteristics of the data and their sources.

The data is limited in many ways. For instance many site specific variables (weather conditions, school quality) that could be relevant to the wage or rent equations are not available for the researcher or are prohibitively expensive. The exclusion of these variables will bias the estimation of the parameters; however it will be difficult to sign the effect on welfare measures since a great level of information would be required in order to have an *a priori* expectation of this effect. I will extend my comments on this issue in the discussion of the econometrics results.

5 Estimation

In this section I present the results of the estimations. In the first part I show the results from the wage and rent equations as far as welfare measures and variances without considering endogeneity or any other estimation issue. The main result of the paper can be derived from these results. In the Second part I discuss some aspects of the estimation that can help to evaluate the results.

5.1 Econometric results

Per the discussion in the econometric section I estimate the model in three ways. The first two estimations use Heckman's two step model. That is, in the first step I estimate an equation of participation. From this equation I calculate the inverse Mill's ratio (lambda) which is *a posteriori* used in the estimation of the wage equations. The difference between model 1 and 2 is that the first model estimates the wage and rent equations separately without any consideration of the simultaneity of the error terms. In contrast, model 2 estimates are obtained using a SURE regression model. Finally, the third and last model is estimated using ML (Tobit system given in equations 7-12). Notice that in the last model the decision of participation is estimated jointly with the wage equation.

As a starting point several location characteristics are incorporated in the estimations, and Table 3 in Appendix 1 presents the most remarkable results. Both PM10 and CRIME have negative signs in the wage equation. Both results are very sensitive to the specification of the wage equation, and I therefore check for multicollinearity. The regressions of PM10 and CRIME on other location characteristics show that the problem is severe. The R-squared values were higher than those in the original wage equation (see Tables 4 and 5 in Appendix 1). Elimination of highly correlated variables changes the unexpected results for PM10 in all the equations, but nevertheless the sign for the crime parameter is not as expected in the OLS wage equations (see

Table 1). I report the model with the more coherent results, that is, the model with significant parameters and the more appropriate sign.

Table 1 presents the outcomes of the estimations of the wage equation. Table 2 in Appendix 1 presents the estimation of the preliminary probit model required by this procedure. The first and third column of Table 1 present the OLS and SURE estimators of the wage equation, including controls for job characteristics. Columns two and four present the same models but with fewer explanatory variables, as I exclude the entire set of variables that cannot be included in the Tobit model, i.e. job characteristics.

The results of this estimation are quite interesting, and most parameters show a rather high level of significance. All parameters associated with the Mincer model of education are significant and have the expected sign. Note that the highly significant coefficient of the gender variable indicates the advantage of being a male worker. Moreover, working in micro-enterprises has negative effects on earning while those in large firms show higher earnings compared to workers in middle-sized firms. The parameter associated with Social Security is positive in one equation and negative in the second, but is not significant in either. As to the employment sectors (for which the baseline is "industry") note that the only negative and significant parameter is that of Commerce (branch 6). Several of these parameters could not be estimated with a reasonable level of statistical significance. Strong and significant effects were observed for Mining (branch 2) and for Financial Services (branch 8), two sectors that traditionally have been associated with higher wages. The inclusion of job characteristics improves the goodness of fit of the model.[2]

The lambda parameter, indicating the sample selection effect, is positive, relatively large and has a rather high level of statistical significance. The parameter corresponding to the environmental variable, airborne particulate matter, is estimated with a high level of significance and is positive as expected from Roback's model. The variable measuring crime is negative in the two first models and significant only in the first OLS model. Curiously the crime parameter changes sign when using a simultaneous equation model. I discuss this issue below. Finally, the unemployment rate is negative and significant in each model.

Table 2 presents the results of estimations of the rent equation. All the parameters are estimated with a rather high level of statistical significance. Note that the environmental variables, both airborne particulate matter and crime, are estimated with a high level of statistical significance. The signs of CRIME, PM10 and UR fit with the structure of the theoretical model.

---

[2] Remember that both OLS and SURE models have been corrected for selection bias.

Table 1. wage equation

	OLS1	OLS2	SURE1	SURE2	TOBIT
Constant	3.3260*	3.1795*	1.9435*	1.9631*	-13.485*
	0.2572	0.2741	0.3103	0.3170	0.5685
AGE	0.0611*	0.0734*	0.1051*	0.1043*	0.6400*
	0.0074	0.0086	0.0098	0.0100	0.0237
AGE2	-0.0006*	-0.0007*	-0.0012*	-0.0012*	-0.0079*
	0.0001	0.0001	0.0001	0.0001	0.0003
EDUC	0.1253*	0.1331*	0.1325*	0.1429*	0.2243*
	0.0029	0.0027	0.0032	0.0031	0.0126
HEAD	0.4390*	0.4558*	0.5841*	0.6308*	2.5020*
	0.0347	0.0384	0.0442	0.0451	0.1575
SEX	0.2790*	0.3446*	0.4907*	0.4936*	3.1537*
	0.0514	0.0457	0.0520	0.0525	0.1429
MICRO	-0.1116*	-	-0.0801*	-	-
	0.0191		0.0247		
LARGE	0.1462*	-	0.1472*	-	-
	0.0167		0.0208		
TEN	0.0945*	-	0.0824*	-	-
	0.0228		0.0272		
SS	0.0402	-	-0.0154	-	-
	0.0248		0.0300		
CONT	0.0390	-	0.0710*	-	-
	0.0234		0.0275		
Branch1	-0.0594	-	-0.0736	-	-
	0.0401		0.0488		
Branch2	0.2999*	-	0.2613*	-	-
	0.0393		0.0501		
Branch4	0.2739*	-	0.1494	-	-
	0.0790		0.0910		
Branch5	0.1040*	-	0.0934	-	-
	0.0337		0.0408		
Branch6	-0.1095*	-	-0.1289*	-	-
	0.0245		0.0317		
Branch7	0.0026	-	-0.0068	-	-
	0.0315		0.0390		
Branch8	0.2174*	-	0.1259*	-	-
	0.0303		0.0359		
Branch9	-0.0042	-	-0.0349	-	-
	0.0224		0.0285		
LAMBDA	0.6251*	0.6118*	1.1150*	1.0458*	
	0.1046	0.1061	0.1192	0.1223	
PM10	0.0016*	0.0017*	0.0025*	0.0026*	0.0062*
	0.0003	0.0004	0.0005	0.0005	0.0026
CRIME	-0.0006*	-0.0001	0.0002	0.0001	0.0012
	0.0002	0.0002	0.0003	0.0003	0.0017
UR	-0.0323*	-0.0366*	-0.0302*	-0.0318*	-0.1113*
	0.0021	0.0022	0.0027	0.0027	0.0148
N	7880	7880	4848	4848	12999
R-square	0.4734	.3625	-	-	-
Log-L	-9789.65	-15666.21	-4528.29	-4651.53	-11060.8
Rest.(b=0) Log-L	-12316.21	-1.80E+04	380	380	-
F	321.02	664.30	-	-	-

* significance at a α=0.05 level. Below the parameters I report the robust standard errors.

Table 2. rent equation

Variable	OLS	SURE1	SURE2	TOBIT
Constant	7.8633*	7.6658*	7.6772*	7.5276*
	0.0683	0.0960	0.0960	0.0366
NUMPER	-0.2900*	-0.2708*	-0.2707*	-0.2711*
	0.0033	0.0040	0.0040	0.0020
ROOM1	0.1526*	0.1482*	0.1484*	0.1533*
	0.0040	0.0044	0.0044	0.0027
EDUC	0.0336*	0.0421*	0.0421*	0.0335*
	0.0015	0.0020	0.0020	0.0012
WC	0.4087*	0.4246*	0.4176*	0.3762*
	0.0356	0.0466	0.0466	0.0167
FLOOR	0.2306*	0.1731*	0.1707*	0.2771*
	0.0396	0.0552	0.0552	0.0231
ROOF	0.2222*	0.3403*	0.3366*	0.2384*
	0.0563	0.0789	0.0789	0.0282
WALL	0.4225*	0.3508*	0.3490*	0.3275*
	0.0293	0.0443	0.0443	0.0199
INCOMEH	8.443E-08*	8.15E-08*	8.141E-08*	6.674E-08*
	1.92678E-08	8.6825E-09	8.6811E-09	1.1919E-09
PM10	0.0033*	0.0031*	0.0031*	0.0035*
	0.0002	0.0004	0.0004	0.0003
CRIME	0.0019*	0.0013*	0.0013*	0.0026*
	0.0001	0.0003	0.0003	0.0002
UR	-0.0365*	-0.0335*	-0.0335*	-0.0174*
	0.0016	0.0022	0.0022	0.0015
N	12999	4848	4848	12999
R-square	0.6766	-	-	-
F (10,12987)	2469.63	-	-	-
Log-L	-14378.82	-3554.38	-3554.51	-11060.8
Rest.(b=0) Log-L	-21715.1141	-447.7241	-477.7241	-

* significance at a α=0.05 level. Below the parameters I report the robust standard errors

Using the results in Tables 1 and 2, I estimate the marginal value of PM10 and CRIME using equation (16). The mean monthly rent is C$8109.98 and the mean level of monthly earnings is C$209276[3]. Table 3 presents the marginal value of PM10 and

---

[3] 1 dollar in 1998 was around 450 pesos.

its standard deviation; in the OLS cases I assume zero covariance, since no covariance can be obtained from the econometric estimation.

In all cases the marginal value is negative as we expect from a disamenity. Since the covariance in the last three models is positive, OLS tends to underestimate the welfare measure and overestimate its variance. The inclusion of variables associated with job characteristics decreases the marginal value of the amenity. Finally, ML Tobit model estimation tends to increase the value of the amenity significantly.

Table 3. Marginal value of PM10 (pesos)

	OLS1	**OLS2**	**SURE1**	**SURE2**	**TOBIT**
PM10: rent equation	0.00328	0.00328	0.00312	0.00313	0.00351
variance	5.848E-08	5.848E-08	1.594E-07	1.594E-07	8.399E-08
PM10: wage equation	0.0016	0.0017	0.0025	0.0026	0.0062
variance	1.203E-07	1.458E-07	2.436E-07	2.475E-07	6.6045E-06
Covariance	0	0	3.26E-08	3.34E-08	9.84E-09
Marginal Value	**-307.31**	**-337.82**	**-488.11**	**-510.26**	**-1264.68**
standard deviation	72.61	79.93	103.34	104.17	537.83
t	-4.23	-4.23	-4.72	-4.90	-2.35

Table 4 presents the marginal value for crimes. Ignoring for a moment the t values, results with respect to crime are very interesting. All OLS estimations suggest that crime is an amenity, however the simultaneous equations suggest the opposite. This comes from different signs on the crime parameters in the wage equation. A reasonable analysis suggests that crime should not be an amenity, and in this sense the ML estimator and SURE estimator are capturing this expectation; however, none of them is significant. In the next section I give some comments on all the results and specifically on the case of crime.

Table 4 Marginal Value of CRIME (pesos)

	OLS	OLS	SURE1	SURE2	TOBIT
CRIME: rent equation	0.0019	0.0019	0.0013	0.0013	0.0026
variance	2.016E-08	2.016E-08	6.897E-08	6.8982E-08	4.11376E-08
CRIME: wage equation	-0.0006	-0.0001	0.0002	0.0001	0.0012
variance	3.807E-08	5.262E-08	1.048E-07	1.093E-07	2.97657E-06
Covariance	0	0	1.57E-08	1.47E-08	1.92E-09
Marginal Value	**137.50**	**39.36**	**-31.77**	**-6.80**	**-220.78**
standard deviation	40.85	48.02	67.80	69.25	361.06
t	3.37	0.82	-0.47	-0.10	-0.61

Finally, the estimation of welfare measures and their standard deviations show us that the selection of a method of estimation is not innocuous as OLS tends to bias both the marginal value and the variance of the welfare measure.

5.2 Comments on the estimations

5.2.1 Endogeneity A common problem within this type of model is the possible endogeneity of explanatory variables. I focus my analysis on the amenities (CRIME, PM10 and UR) since we are more interested in the parameters associated with these variables. My theoretical model does not explicitly consider the endogeneity of any of these variables; therefore I have to deal with the issue only from a statistical point of view. In this sense I will used the definition of endogeneity as $E(x_i\varepsilon_i) \neq 0$, where x_i is any explanatory variable and ε_i is the error term in a regression.

In order to address the problem of endogeneity I test the hypothesis that the variable x_i is exogenous using a Hausman test. The basic idea can be explained as follows (Wooldrige 2002). The population model is

$$y_1 = \beta x + \alpha y_2 + \varepsilon_1$$

where x includes the exogenous variables in the population equation, and y_2 is a variable whose exogeneity is being tested. We can group all the exogenous variables with a set of instruments called $\tilde{z}$, in $z = (\tilde{z}, x)$. We assume that $E(z'\varepsilon_1) = 0$. Let's write the linear projection of y_2 on z as

$$y_2 = \pi z + \upsilon \tag{18}$$

with $E(z'\upsilon) = 0$. Define $\varepsilon_1 = \rho\upsilon + error$. Then the population equation can be written as

$$y_1 = \beta x + \alpha y_2 + \rho\upsilon + error$$

A test of exogeneity could be expressed as

$$H : \rho = 0$$

Even though we do not know υ we can use the OLS estimator given by $\hat{\upsilon} = y - \hat{\pi} z$ and estimate the model

$$y_1 = \beta x + \alpha y_2 + \rho\hat{\upsilon} + error$$

I follow this procedure to test the exogeneity of Crime, PM10 and UR. The null hypothesis is rejected at a 1% for the last two variables in both the wage and rent equations (I report the results for the wage equation only: see Tables 6 and 7 for PM10, 8 and 9 for Crime and 10 and 11 for UR, all in Appendix 1). The instrumental variables that I use are some of the site specific variables that are highly correlated with the variables under analysis. It is not clear that these are the best instruments, but they are the only information available so far.

The models change in the following directions. Since PM10 and UR are endogenous according to the Hausman test, the OLS changes to a two-step least square method (2SLS). Each equation, wage and rent, is estimated separately using 2SLS. The SURE model changes to a three-step least square method (3SLS); that is, I simultaneously estimate the rent and wage equations using instrumental variables. Finally for the Tobit method I use a similar procedure for the 3SLS estimator. First, I estimate equation 18 and then I use the predicted value ($\hat{y}_2$) in the estimation of the simultaneous Tobit model.

The following tables show results analogous to those presented in Tables 1 and 2, where I do not consider variables associated with the job characteristics.

Table 5. Wage Equation

Variable	2SLS	t	3SLS	t	TOBIT	t
Constant	3.0848	10.23	3.7112	6.06	-16.1610	-20.89
AGE	0.0744	8.05	0.1105	8.87	0.6517	23.12
AGE2	-0.0008	-6.79	-0.0013	-8.14	-0.0080	-24.32
EDUC	0.1298	42.97	0.1088	20.55	0.2171	14.22
HEAD	0.4660	11.35	0.5687	9.10	2.5286	13.75
SEX	0.3160	6.16	0.5763	7.41	3.2165	19.05
LAMBDA	0.5898	5.02	1.0371	6.83	-	-
PM10	0.0037	8.03	0.0135	2.27	0.0254	6.26
CRIME	-0.0004	-2.04	0.0033	2.21	0.0068	3.38
UR	-0.0411	-16.58	-0.2962	-13.81	-0.0545	-2.69
N	5971.0		4848.0		9804.0	
Log_L	-7435.6		-7273.8		-7756.0	
Rest.(b=0) Log-L	-9113.0		386.7		-	

Table 6. Rent Equation

Variable	2SLS	t	3SLS	t	TOBIT	t
Constant	8.0054	105.60	8.9123	5.16	7.5750	173.57
NUMPER	-0.2850	-85.85	-0.1644	-2.94	-0.2625	-112.45
ROOM1	0.1344	35.01	0.1083	2.33	0.1331	42.75
EDUC	0.0278	20.32	0.0174	1.15	0.0240	17.77
WC	0.4096	9.05	0.3050	0.67	0.3882	21.34
FLOOR	0.2722	6.33	-0.2290	-0.80	0.3144	14.14
ROOF	0.2903	4.69	0.1774	0.75	0.2371	8.81
WALL	0.4080	12.57	-0.1315	-0.30	0.3378	16.64
INCOMEH	0.0000	9.16	0.0000	-2.81	0.0000	50.30
PM10	0.0034	11.37	0.0625	2.67	0.0037	17.27
CRIME	0.0020	13.05	0.0157	2.73	0.0059	13.34
UR	-0.0479	-27.79	-0.5821	-2.33	-0.0372	-18.03
N	9804		4848		9804.0	
Log_L	-10490.69		-10194.30		-7756.0	
Rest.(b=0) Log-L	-16117.62		-477.72			

In general the parameters of PM10 and CRIME tend to be greater in these equations than in those presented in Tables 1 and 2, but signs of the coefficients are the same. The welfare measures are given in the following tables:

Table 7. Marginal Value of PM10 (pesos)

	2SLS	3SLS	TOBIT
Rent equation parameter	3.44E-03	6.25E-02	5.86E-03
variance	9.15E-08	5.50E-04	1.93E-07
Wage equation parameter	3.72E-03	1.35E-02	2.54E-02
variance	2.14E-07	3.57E-05	1.65E-05
Covariance	0.00	-7.19E-06	-4.23E-08
Marginal Value	**-750.31**	**-2326.36**	**-5273.45**
standar deviation	96.89	1113.23	849.49
t-value	-7.74	-2.09	-6.21

Table 8. Marginal Value of CRIME (pesos)

	2SLS	3SLS	TOBIT
Rent equation parameter	1.97E-03	1.57E-02	3.75E-03
variance	2.27E-08	3.28E-05	4.70E-08
Wage equation parameter	-4.32E-04	3.25E-03	6.77E-03
variance	4.50E-08	2.17E-06	4.00E-06
Covariance	0.00	-2.14E-07	-6.66E-09
Marginal Value	**106.3**	**-554.1**	**-1385.7**
SD	44.42	275.28	418.67
t	2.39	-2.01	-3.31

The absolute values of the welfare measure change considerably when we consider the endogeneity of UR and PM10, but the main conclusion of the paper holds. That is, it is relevant to consider the simultaneity of the wage and rent equations in the hedonic wage model because doing so has a significant effect on the welfare measures.

5.3 Crime disaggregation

Additional analysis can be done in the case of crime. Following Deller et al. (2001), I disaggregate the crime data into different types of crime and reestimate an OLS model. Results are presented in Tables 12 and 13 in Appendix 1. Similar to Deller's results, different types of crime have different effects in both wage and rent equations. Accordingly, a different analysis must be done taking into account the different types of crime in order to obtain a welfare measure compatible with the category of crime used. Notice that the coefficients of some crime categories are positive and other are negative. Furthermore, some are significant and others are not. This complexity is not captured by OLS, because in the first model we have the wrong sign and a significant parameter. I do not report welfare measures for type of crime since my objectives are mostly achieved with results in Tables 3 and 4.

5.4 Omitted variables

Specification of the hedonic model is crucial for the correct interpretation of the welfare measures. Any omitted variable will bias the result. It is not clear that the available data considers all the relevant explanatory variables. Some commonly used variables in other studies such as weather and school quality were not available for this research.

In order to define the direction of this bias we would need to know the sign of the parameters associated with the explanatory variable quality of education (or other) in both wage and rent equations. The traditional equation for the bias in the omitted variable case is given by

$$\hat{\beta}_k = \beta_k + \gamma\delta_k$$

where the estimated parameter $\hat{\beta}_k$ is equal to the true parameter β_k plus the bias $\gamma\delta_k$. This bias is given by the parameter of the omitted variable in the principal equation γ (wage or rent in this case) and the parameter δ_k which is the coefficient in an auxiliary regression of the variable under interest x_k on the omitted variable. In this case these regressions will be crime and PM10 on quality of education (or other).

Following the analysis in section 2 and taking quality of education as an amenity, the sign of γ on the wage equation should be negative and ambiguous in the rent equation. On the other hand, we do not have a prior for the sign of δ_k, the quality of education in the crime or PM10 equations. Therefore, the direction of the effect of omitting this variable on the parameter $\hat{\beta}_k$ is not clear, so we cannot suggest any effect on the welfare measures. This is a clear weakness of the econometric estimation, which is an important consideration if the welfare measures are used for cost benefit analysis. Notice however that this is a very general problem in applied econometrics: we must be aware of this issue and consider its implications, but there is nothing else we can do if we do not have a full characterized model, which requires that we know all the relevant variables and perform our estimation using optimal data. Unfortunately, this data is not available, and a perfect model is not estimated. Fortunately, the first objective of this paper can still be achieved, and the conclusions are not changed: simultaneity of the wage and rent equations should be considered when estimating a hedonic wage model.

5.5 Welfare measures

My last comments about the results are related to the use of welfare measures. Most of the aforementioned literature has used marginal values as weights for quality of life indices. I do not follow this approach since from my point of view, welfare measures should be a direct input for policy makers as they provide the basis for the cost benefit analysis of public policies that involve changes in the level of these amenities. The marginal price times the number of workers in a community can be a rough estimation of the marginal value of a change in the level of the amenity. For instance considering only employed individuals from the Chilean capital, the value of a marginal change in air quality will be in the range of 0.02 to 0.09 percent of the GDP for Santiago's Metropolitan Region. These results are intuitively reasonable. Moreover, this estimation compares favorably with other estimations done using the hedonic price approach such as Rogat and Firenguetti (1996). Valuing a non-marginal change requires a general equilibrium model like those proposed in Giannais (1988) and Sieg et al. (2003), which is clearly beyond the scope of this paper given the available data.

Furthermore, another approach known as the two step method suggested by Rosen (1979) and used in some papers such as Kahn (1988), has been discredited because it does not provide useful information for aggregation (see Brown and Rosen, 1982).

6 Conclusion

Air quality and crime seem to affect the differentiation of wages in Chile. The marginal value for air pollution is significant and ranges between -307 and -5273 pesos depending of the methods of estimation. Using an aggregate measure of crime the models give the value of crime as between -1385 and 106 pesos, which are mostly not significant. This result shows that researchers must be aware of the different effects of the different types of crime. Further analysis suggests that the decomposition of crime into different types produces better and more interesting results.

Considerable attention should be paid to the endogeneity of the amenities when the welfare measures estimated in the model will be used in a cost benefit analysis. My results suggest that ignoring the endogeniety tends to result in downward bias in the welfare measure.

The main conclusion of the paper, which holds regardless of endogeneity, is that the selection of an econometric estimation method that does not consider the simultaneity of the wage and rent equation has considerable effects on the estimated values and variance of welfare measure. The results show that OLS (2SLS) underestimates welfare effects and their variance. Additionally, with more complicated amenities such as crime, OLS (2SLS) tends to obtain the wrong sign for the parameters in the wage equation. On the other hand simultaneous estimation using GLS (3SLS) underestimates the welfare measure in comparison with ML, although both methods capture correctly the sign and the significance of the parameters associated with the amenities.

Some possible flaws of the results and possible extensions for the research are the following: the econometric estimation does not consider the cluster effects on the variance of parameters which should increase the variance of the OLS estimator for clustered variables and thereby decrease the t values. A correct evaluation of this problem requires an understanding of this phenomenon in all three models (OLS heckit corrected, SURE and ML).

There is always room for improvement in the quality of data, especially in the environmental quality and crime information. Other municipalities should be incorporated in data as soon as environmental data is available. More environmental quality variables and other non-environmental amenities should be incorporated in the analysis as well. Gains on information could be achieved from a disaggregate analysis of crime similar to the analysis presented in section 5.2.

7 References

1. Anderson, R. and T. Crocker (1971), "Air Pollution and Residential Property Values" *Urban Studies*, **vol. 8** Issue 3 pp. 171-80.

2. Anderson, R. and T. Crocker (1972), "Air Pollution and Property Values: A Reply" *Review of Economics and Statistics*, **vol. 54** Issue 4 pp. 470-73.

3. Bartik, T. and K. Smith, (1987), "Urban Amenities and Public Policy" in "*Handbook of Regional and Urban Economics*", E. Mills, editor, Amsterdam, The Netherlands: Elsevier.

4. Bender, B., T. Gronberg and H. Hwang, (1980), "Choice of Functional Form and the Demand for Air quality" *Review of Economics and Statistics*, **vol. 62** Issue 4 pp. 638-43.

5. Berger, M., G. Blomquist and K. Sabiriniava, (2003) "Compensating Differentials in Emerging Labor and Housing Markets; Estimates of Quality of life in Russian Cities" IZA discussion papers series, Institute for the Study of labor. pp. 53.

6. Blomquist, G., M. Berger and J.Hoehn, (1988) "New estimates of Quality of Life in Urban Areas" *The American Economic Review*, **vol. 78**, pp. 89-107.

7. Brown, J. and S. Rosen (1982), "On Estimation of Structural Hedonic Price

Models" *Econometrica*, **vol. 50**, Issue 3, pp. 765-68.

8. Cassel, E. and R. Mendelsohn, (1985) "The Choice of Functional forms in Hedonic Price Equations: A Comment" *Journal of Urban Economics,* **vol. 18**, pp. 135-42.
9. Clark, D. and J. Kahn (1988), "The Two-Stage Hedonic Wage Approach: A Methodology for Valuation of Environmental Amenities" *Journal of Environmental Economics and Management,* **vol. 16**, pp. 106-20.
10. Cropper, M., B. Deck and K. McConnell, (1988), "On the Choice of Functional Form for Hedonic Price Functions", *Review of Economics and Statistics*, **vol. 70** Issue 4 pp. 668-75.
11. Deller, S. and T. Ottem (2001), "Crime and the Quality of Life in Wisconsin Counties" *Agricultural and Applied Economics Staff Papers*, University of Wisconsin. pp.31.
12. Figueroa, E. and G. Lever, (1992), "Determinantes del Precio de la Vivienda en Santiago: Una Estimación Hedónica", *Estudios de Economía*, **vol. 19**, Issue 1 pp. 69-83.
13. Figueroa, E., J. Rogat and L. Firinguetti (1996), "An Estimation of the Economic Value of Air Quality Improvement Program in Santiago de Chile " Es-

tudios de Economía, **vol. 23**, special number August. pp. 101-14.

14. Freeman, M., (1971), "Air Pollution and Property Values: A Methodological Comment" *Review of Economics and Statistics*, **vol. 53** Issue 4 pp. 415-16.
15. Freeman, M., (1974), "Air Pollution and Property Values: A Further Comment" *Review of Economics and Statistics*, **vol. 56** Issue 4 pp. 554-56.
16. Freeman, M., (1979), "The Benefits of Environmental Improvement: Theory and Practice", Baltimore: John Hopkins University Press.
17. Freeman, M., (2003), The Measurement of Environmental and Resource Values. Theory and Methods, Washington D.C.: Resources for the Future
18. Graves, P., C. Murdoch, M. Thayer and D. Walmand, (1988), "The Robustness of Hedonic Price Estimation: Urban Air Quality", *Land Economics,* **vol. 64**, pp.220-33.
19. Heckman, J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, **vol. 47**, Issue 1, pp. 153-61.
20. Hoehn, J., M. Berger and Glenn C. Blomquist (1988) "A Hedonic model of Interregional Wages, Rents and Amenity Values" *Journal of Regional Sciences,* **vol. 27**, Issue 4, pp. 605-620.

21. Maddala, G. (1983) Limited Dependent and Qualitative variables in Econometrics, Cambrige University press.

22. Marin, A. and Psacharopoulos, G., (1982), "The Reward for Risk in the Labor Market . Evidence from the United Kingdom and a Reconcilation with other Studies", *Journal of Political Economy*, **vol. 90**, Issue 4, pp. 827-53.

23. McConnell, K., (1995), "Issues in Estimating Benefits with Non-Market Methods", Washington D.C.: Inter-American Development Bank.

24. Nelson, F. and L. Olson (1978) "Specification and Estimation of a Simultaneous-Equation Model with Limited Dependent Variables" *International Economic Review,* **vol. 19**, pp. 695-709.

25. Palmquist, R., and L. Danielson, (1989), "A Hedonic Study of The effects of Erosion Control and Drainage on Farmland Values", *American Journal of Agricultural Economics*, **vol. 71** pp. 55-62.

26. Palmquist, R., (1991): "Hedonic Methods" in J.Braden & C. Kolstad, editors, Measuring the Demand for Environmental Improvement, Amsterdam, The Netherlands: Elsevier.

27. Palmquist, R., F. Roka and T. Vukina, (1997), "Hog Operations, Environmental Effects, and Residential Property Values", *Land Economics*, **vol 73**, Issue 1 pp:

114-24.

28. Ridker, R. and Henning (1967), "The Determinants of Residential Property Values with Special Reference to Air Pollution". *Review of Economics and Statistics*, **vol. 49**, Issue 2 pp.247-57.

29. Roback, J., (1982), "Wages, Rents, and the Quality of Life", *The Journal of Political Economy*, **vol. 90**, Issue 6 (Dec.), pp. 1257-1278.

30. Roback, J., (1988), "Wages, Rents, and Amenities: Differences Among Workers and Regions", *Economic Enquiry,* **vol. 26**, pp. 23-41.

31. Rosen, S (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition", *Journal of Political Economy*, **vol. 82**, Issue 1, pp.34-55.

32. Rosen, S. (1979): "Wage-Based Indices of Urban Quality of Life" in P.Mieszkowski & M.Straszheim, editors, Current Issues in Urban Economics, Baltimore, John Hopkins University Press.

33. Rosen, S (1986): "The Theory of Equalizing Differences" in O.Ashenfelter & R Layard, editors, *Handbook of Labor Economics*, **vol.1**, Amsterdam, Noth-Holland, pp. 641-92.

34. Smith and Deyak (1975), "Measuring the Impact of Air Pollution on Property Values" *Journal of Regional Science*, **vol. 15**, Issue 3, pp. 277-88.

35. Wooldrige, Jeffrey (2002) Econometric Analysis of Cross Section ands Panal Data, The Mit Press, Cambridge, Massachusetts.

36. Wunder, D. & H. Gutierrez (1992), “Modelos Tipo Box-Cox: Proposición de una Estrategia de Estimación” *Revista de Análisis Económico*, **vol. 7**, Issue 2 pp. 79-102.

8 Appendix 1

Table 1. Descriptive Statistics

Variable	source	Mean	Std.Dev.	Minimum	Maximum	Cases
Age (AGE)	CASEN	40.56	16.45	18	99	19584
Education (EDUC)	CASEN	9.87	4.4	0	21	19436
Marital Status (MS) *	CASEN	0.61	0.49	0	1	19584
Household's Head (HEAD)*	CASEN	0.39	0.49	0	1	19584
Sex (SEX)*	CASEN	0.48	0.5	0	1	19584
Rents (RENT)	CASEN	29170.81	36860.88	0	439000	19584
Log of rents (LRENT)	CASEN	6.94	4.93	0	12.99	19584
Percapita rent (RENTCAP)	CASEN	8109.98	12744.57	0	175600	19584
Log of percapita rent (LRENTCAP)	CASEN	5.99	4.29	0	12.08	19584
log of wage (LNY)	CASEN	6.85	0.9	2.97	11.75	10598
Participation in Labor Market (PART)*	CASEN	0.55	0.5	0	1	19584
Micro Enterprise (MICRO)*	CASEN	0.44	0.5	0	1	10838
Large enterprise (LARGE)*	CASEN	0.22	0.41	0	1	10838
Tenure (TEN)*	CASEN	0.75	0.43	0	1	10838
Social Security (SS)*	CASEN	0.69	0.46	0	1	19584
Contract (CONT)*	CASEN	0.78	0.41	0	1	7988
Branch 1 Agriculture and Fishing (B1)*	CASEN	0.14	0.35	0	1	10838
Branch 2 Mining (B2)*	CASEN	0.04	0.19	0	1	10838
Branch 4 Electricity, Gas and Water (B3)*	CASEN	0.01	0.11	0	1	10838
Branch 5 Building (B4)*	CASEN	0.08	0.27	0	1	10838
Branch 6 Commerce B(5)*	CASEN	0.18	0.38	0	1	10838
Branch 7 Transp., Storage and Com. (B6)*	CASEN	0.08	0.27	0	1	10838
Branch 8 Financial service (B7)*	CASEN	0.06	0.24	0	1	10838
Branch 9 Social and Com. services (B8)*	CASEN	0.27	0.44	0	1	10838
Number of rooms (ROOMS)	CASEN	4.43	2.01	1	16	19584
Household's size (NUMPER)	CASEN	6.07	2.15	0	26	19584
Water Close (WC)*	CASEN	0.85	0.35	0	1	19584
type of Floor (FLOOR)*	CASEN	0.95	0.22	0	1	19584
type of Roof (ROOF)*	CASEN	0.97	0.18	0	1	19584
type of wall (WALL)*	CASEN	0.92	0.27	0	1	19584
Household's Income (INCOMEH)	CASEN	554154	991330	0	41797900	19584
Particulate Material (PM10)	CONAMA/SESMA	68.86	21.36	16	105.58	19584
CRIME/'000.000	INE/POLICE	52.13	30.61	20.11	158.52	19584
number of Vehicles (V1)	INE	27288.39	21757.52	1200	84038	18901
unemployment rate (UR)	INE	8.93	3.59	2.9	17.4	19584
Percapita budget (PB)	SINIM	73.43	57.6	31.2	302.7	19584
poverty rate (POOR)	SINIM	17.29	8.26	1.4	32	14965
population density (DENS)	SINIM	2201.63	3538.51	1.22	13540.85	18201
Potable water coverage (COVWAT)	SINIM	0.85	0.11	0.6	1	19584
energy power coverage (COVELEC)	SINIM	0.7	0.16	0.3	0.9	19584
Cleaning services expenditures (SERV)	SINIM	0.37	0.16	0	0.7	19584
Water expenditures (WATER)	SINIM	0.03	0.05	0	0.2	19584
Public lighting expenditures (ALUM)	SINIM	0.22	0.13	0.1	0.7	19584
% of Urban area (URB)	SINIM	0.21	0.33	0	1	19584
% of land without vegetation (VEG)	SINIM	0.21	0.36	0	1	19584
Chilean Capital (SAN)*	SINIM	0.29	0.46	0	1	19584

* Dummy variables, CASEN= National Survey on Socioeconomic Characterization, CONAMA=National Committee for the Environment

INE= Institute of National Statistics SINIM= National Municipal Information System

Table 2. Probability of being employed (probit Estimation)

Variable	**Coefficient**	**Standar error**	**t**
Constant	-3.30450	0.08800	-37.55200
AGE	0.14669	0.00412	35.61300
AGE2	-0.00181	0.00005	-39.69200
EDUC	0.03837	0.00255	15.03100
MS	-0.14174	0.02275	-6.23100
HEAD	0.74316	0.02755	26.97700
SEX	0.79053	0.02330	33.93200
N	19436	Degree of freedom	6
Log Likelihood	-10118.45	Chi-squared	6482.38
Rest. Log Likelihood	-13359.64		

Tabla 3. OLS estimation wage equation

Variable	**Coefficient**	**Standar error**	**t**
constant	2.987495655	0.262988	11.36
AGE	7.17E-02	0.007386	9.709
AGE2	-7.61E-04	0.000091	-8.345
EDUC	1.19E-01	0.002972	39.918
HEAD	4.68E-01	0.035311	13.262
SEX	3.62E-01	0.051125	7.073
MICRO	-1.07E-01	0.019408	-5.488
LARGE	0.152824919	0.016845	9.072
TEN	6.75E-02	0.023560	2.864
SS	6.55E-02	0.025182	2.601
CONT	4.79E-02	0.023599	2.032
B1	1.25E-02	0.040681	0.308
B2	0.316511602	0.041296	7.664
B4	2.91E-01	0.080535	3.614
B5	0.103167226	0.033706	3.061
B6	-1.03E-01	0.024719	-4.163
B7	-3.12E-03	0.031782	-0.098
B8	1.72E-01	0.030489	5.65
B9	-1.53E-02	0.022423	-0.683
LAMBDA	7.49E-01	0.105092	7.128
PM10	**-4.60E-03**	**0.000928**	**-4.953**
CRIME	**-3.70E-03**	**0.000320**	**-11.573**
DENS	-1.13E-05	0.000003	-3.357
V1	3.80E-06	0.000001	5.395
UR	-8.75E-03	0.003643	-2.401
PB	1.67E-03	0.000175	9.543
SERV	0.355311823	0.094550	3.758
COVWAT	1.244945342	0.270789	4.597
VEG	4.53E-02	0.050488	0.898
URB	0.23112037	0.068652	3.367
SAN	0.238532625	0.031485	7.576
N	7332	R-square	.503264
Log-L	-8823.3267	Rest. (b=0) Log-L	-11388.414

Table 4. OLS PM10

Variable	Coefficient	s.d.	t
Constant	29.941	0.717	41.743
SAN	19.707	0.297	66.265
POOR	0.244	0.014	16.941
DENS	0.001	0.000	11.465
COVELEC	45.210	0.797	56.744
SERV	16.787	0.620	27.088
WATER	19.893	2.447	8.129
URB	-16.246	0.863	-18.829
VEG	-22.010	0.353	-62.385
N	8045	R-square	0.84
Log-L	-27447.97	Rest. (b=0) Log-L	-34766.91

Table 5. OLS CRIME

Variable	Coefficient	s.d.	t
Constant	137.886	2.787	49.472
SAN	-36.663	0.666	-55.040
V1	0.000	0.000	12.049
DESP	1.854	0.092	20.233
P10	0.233	0.008	30.353
POOR	0.333	0.055	6.019
DENS	-0.006	0.000	-23.761
COBERA	-341.231	6.617	-51.573
COVELEC	159.413	5.432	29.347
SERV	51.273	2.061	24.878
WATER	59.364	5.040	11.779
ALUM	26.482	2.386	11.100
URB	134.052	3.266	41.043
VEG	47.467	0.829	57.244
N	8045	R-square	0.81
Log-L	-33096.52	Rest. (b=0) Log-L	-39897.35

Table 6. OLS PM10 (testing endogeneity)

Variable	**Coefficient**	**s.d.**	**t**
Constant	28.652	1.272	22.521
AGE	0.074	0.048	1.543
AGE2	-0.001	0.001	-1.676
EDUC	-0.038	0.026	-1.470
HEAD	-0.203	0.237	-0.857
SEX	0.315	0.227	1.391
SAN	20.034	0.344	58.263
POOR	0.273	0.017	16.018
DENS	0.0009	0.0001	9.69
COVELEC	45.515	0.922	49.388
SERV	15.681	0.714	21.969
WATER	13.220	2.883	4.585
URB	-15.985	0.992	-16.111
VEG	-22.369	0.416	-53.785
N	5971	**R-square**	**0.84**
Log_L	-20358.66	Rest.(b=0) Log-L	-25774.47

Table 7. OLS wage equation (testing endogeneity of PM10)

Variable	**Coefficient**	**s.d.**	**t**
Constant	2.887	0.297	9.724
AGE	0.078	0.009	8.544
AGE2	-0.001	0.000	-7.266
EDUC	0.130	0.003	43.762
HEAD	0.481	0.040	11.921
SEX	0.339	0.050	6.764
LAMBDA	0.640	0.115	5.567
PM10	0.004	0.000	8.932
CRIME	0.000	0.000	-2.061
DESP	-0.036	0.002	-16.117
v	**-0.013**	**0.001**	**-10.660**
N	5971	R-square	0.442
Log_L	-7368.24	Rest.(b=0) Log-L	-9108.04

Table 8. OLS CRIME (testing endogeneity)

Variable	Coefficient	s.d.	t
Constant	128.752	4.523	28.468
AGE	-0.089	0.170	-0.526
AGE2	0.001	0.002	0.291
EDUC	0.135	0.092	1.468
HEAD	1.065	0.842	1.264
SEX	-1.267	0.806	-1.573
SAN	-44.333	1.222	-36.267
POOR	-1.227	0.061	-20.258
DENS	-0.002	0.0003	-6.43
COVELEC	-62.646	3.276	-19.122
SERV	-49.135	2.537	-19.365
WATER	18.875	10.250	1.841
URB	106.101	3.527	30.082
VEG	24.034	1.478	16.256
N	5971	R-square	0.430
Log_L	-27931.86	Rest.(b=0) Log-L	-29612.65

Table 9. OLS wage (testing endogeneity of CRIME)

Variable	Coefficient	s.d.	t
Constant	3.245	0.305	10.653
AGE	0.074	0.009	8.008
AGE2	-0.001	0.000	-6.730
EDUC	0.131	0.003	43.365
HEAD	0.464	0.041	11.366
SEX	0.311	0.051	6.057
LAMBDA	0.587	0.117	5.002
PM10	0.001	0.001	2.897
CRIME	-0.001	0.000	-2.693
DESP	-0.036	0.002	-15.972
v	**0.001**	**0.001**	**1.255**
N	5971	R-square	0.431
Log_L	-7424.85	Rest.(b=0) Log-L	-9108.04

Table 10. OLS UR (testing endogeneity)

Variable	Coefficient	s.d.	t
Constant	-4.807	0.348	-13.799
AGE	0.030	0.013	2.290
AGE2	-0.0003	0.0002	-1.704
EDUC	0.010	0.007	1.459
HEAD	-0.223	0.065	-3.435
SEX	0.184	0.062	2.964
SAN	0.974	0.094	10.345
POOR	0.433	0.005	92.882
DENS	-0.001	0.000	-32.117
COVELEC	11.182	0.252	44.316
SERV	-7.365	0.195	-37.688
WATER	-8.888	0.789	-11.259
URB	10.133	0.272	37.302
VEG	0.036	0.114	0.314
N	5971	R-square	0.679
Log_L	-12623.966	Rest.(b=0) Log-L	-16018.274

Table 11. OLS Wage (testing Endogeneity UR)

Variable	Coefficient	s.d.	t
Constant	3.248	0.301	10.789
AGE	0.075	0.009	8.127
AGE2	-0.001	0.000	-6.849
EDUC	0.131	0.003	43.286
HEAD	0.466	0.041	11.395
SEX	0.316	0.051	6.157
LAMBDA	0.593	0.117	5.059
PM10	0.002	0.000	3.504
CRIME	-0.001	0.000	-3.257
DESP	-0.041	0.003	-16.123
v	***0.017***	***0.005***	***3.633***
N	5971	R-square	0.432
Log_L	-7419.58	Rest.(b=0) Log-L	-9108.04

Table 12. OLS wage equation

Variable	Coefficient	s.d.	t
Constant	2.481	0.503	4.931
AGE	0.095	0.015	6.381
AGE2	-0.001	0.000	-5.811
EDUC	0.132	0.005	25.062
HEAD	0.636	0.073	8.727
SEX	0.430	0.083	5.195
MICRO	-0.112	0.034	-3.319
LARGE	0.138	0.028	4.951
TEN	0.132	0.039	3.409
SS	-0.048	0.044	-1.105
CONT	0.070	0.041	1.711
B1	-0.059	0.073	-0.803
B2	0.204	0.073	2.788
B3	0.265	0.159	1.662
B4	0.102	0.061	1.687
B5	-0.101	0.043	-2.335
B6	0.050	0.053	0.931
B7	0.120	0.051	2.379
B8	-0.005	0.038	-0.137
LAMBDA	1.080	0.208	5.203
UR	-0.030	0.003	-8.789
CRIME1	-0.369	0.061	-6.046
CRIME2	0.034	0.031	1.101
CRIME3	-0.038	0.011	-3.325
CRIME4	-0.016	0.003	-4.583
CRIME5	0.076	0.016	4.763
CRIME6	0.027	0.004	6.085
CRIME7	-0.028	0.019	-1.452
CRIME8	0.334	0.088	3.804
N	2702	R-square	0.55
Log-L	-14378.83	Rest.(b=0) Log-L	-3972.25

Table 13. OLS rent equation

Variable	Coefficient	s.d.	t
Constant	7.986	0.171	46.731
NUMPER	-0.269	0.008	-34.094
ROOM1	0.154	0.006	23.839
EDUC	0.037	0.003	13.043
WC	0.292	0.083	3.495
FLOOR	0.218	0.083	2.638
ROOF	0.243	0.123	1.969
WALL	0.388	0.068	5.733
INCOMEH	0.000	0.000	3.491
DESP	-0.027	0.003	-8.957
CRIME1	-0.393	0.053	-7.464
CRIME2	0.022	0.028	0.786
CRIME3	-0.045	0.010	-4.677
CRIME4	-0.004	0.003	-1.260
CRIME5	0.118	0.014	8.235
CRIME6	0.022	0.004	5.653
CRIME7	-0.043	0.016	-2.718
CRIME8	0.250	0.072	3.493
N	2702	R-square	0.712
Log-L	-2393.394	Rest.(b=0) Log-L	-4074.115

CRIME1= crime against order and public security

CRIME2= crime against people

CRIME3= crime against properties

CRIME4= traffic accidents

CRIME5= crime against special laws

CRIME6= demeanor crimes

CRIME7= other crimes

CRIME8= intra-familiar violence

9 Appendix 2

9.1 Heckman's Model

In a labor market context, earnings are determined by a two step decision process. First an agent decides whether or not to participate in the labor market, and if he decides to participate he then determines the number of hours to supply given the market salary and his socoieconomic characteristics. An agent compares his reserve wage (w_r) with the salary that he can obtain in the market (w_m). If $w_r > w_m$ then the market wage does not cover his opportunity cost and this consumer prefers not to work in the market. The reserve salary is assumed to depend on socioeconomic characteristics as well as individual preferences.

Labor market data provides observations of the dependent variable only for people that have already made their participation decision; market wages paid to actual workers are censored samples containing observations of positive salaries for people with $w_r < w_m$ but only a mass point in zero for people with $w_r > w_m$. Due to this arbitrary truncation using OLS to estimate the model will bias the estimator. Thus the model should be rewritten as a two equation model:

$$w_i = x_i'\beta + \varepsilon_i,$$

and

$$\triangle w = w_{ri} - w_m = z_i'\gamma + \upsilon_i,$$

reflecting the fact that there exists a relationship between the salary earned by an agent and her reserve salary. We construct a a dichotomous proxy variable such that

$$\begin{aligned} I &= 1 \text{ if } \triangle w = z_i'\gamma + \upsilon_i > 0 \\ I &= 0 \text{ otherwise.} \end{aligned}$$

Under this structure we know that we observe w_i only if $I = 1$. In order to go further with the econometric estimation some structure must be assumed for the equations. Following Maddala (1983) $E(\varepsilon_i|x,z) = \mu_\varepsilon = 0$, $V(\varepsilon_i|x,z) = \sigma_\varepsilon^2$, $E(\upsilon_i|x,z) = \mu_\upsilon$, $V(\upsilon_i|x,z) = \sigma_\upsilon^2$ and $cov(\varepsilon_i,\upsilon_i|x,z) = \sigma_{\varepsilon\upsilon} = \rho\sigma_\varepsilon\sigma_\upsilon$. with ρ the correlation coeficient between ε_i υ_i.Then the conditional mean of the wage equation is

$$E\left(w_i|I = 1\right) = x_i'\beta + E(\varepsilon_i|I = 1) = x_i'\beta + E(\varepsilon_i|z_i'\gamma + \upsilon_i > 0), \tag{19}$$

the last term in equation 19 shows that the expected OLS estimators will be biased, due to the omitted variables. Assuming that the errors have a joint normal distribution and the conditional expectation of ε_i given υ_i is linear in υ_i we can conclude that .[4] From this distribution we can conclude that

$$\begin{aligned} E(\varepsilon_i|\upsilon_i) &= \mu_\varepsilon + \frac{\sigma_{\varepsilon\upsilon}}{\sigma_\upsilon^2}\left(\upsilon_i - \mu_\upsilon\right) = \frac{\sigma_{\varepsilon\upsilon}}{\sigma_\upsilon^2}\left(\upsilon_i - \mu_\upsilon\right) = \frac{\sigma_{\varepsilon\upsilon}}{\sigma_\upsilon^2}\left(\upsilon_i - \mu_\upsilon\right), \\ &= \frac{\rho\sigma_\varepsilon\sigma_\upsilon}{\sigma_\upsilon^2}\left(\upsilon_i - \mu_\upsilon\right) = \frac{\rho\sigma_\varepsilon}{\sigma_\upsilon}\left(\upsilon_i - \mu_\upsilon\right), \end{aligned}$$

and

$$Var(\varepsilon_i|\upsilon_i) = \sigma_\varepsilon^2(1-\rho^2).$$

---

[4] Hogg and Craig (1978), Introduction to mathematical statistics. pp. 74-76.

Using this results in 19 we have

$$\begin{aligned} E\left(w_i|I=1\right) &= x_i'\beta + E(\varepsilon_i|\upsilon_i > -z_i'\gamma) = x_i'\beta + \frac{\rho\sigma_\varepsilon}{\sigma_\upsilon}\left(E(\upsilon_i|\upsilon_i > -z_i'\gamma) - \mu_\upsilon\right), \\ Var\left(w_i|I=1\right) &= var(\varepsilon_i|\upsilon_i > -z_i'\gamma) = +\frac{\rho^2\sigma_\varepsilon^2}{\sigma_\upsilon^2}Var(\upsilon_i|\upsilon_i > -z_i'\gamma) + \sigma_\varepsilon^2(1-\rho^2). \end{aligned}$$

The last step in order to attain a complete characterization of the statistical basis of the model is to define the $E(\upsilon_i|\upsilon_i > -z_i'\gamma)$ and $Var(\upsilon_i|\upsilon_i > -z_i'\gamma)$. These two components can be derived using the moment generatong function of a truncated normal distribution $f(\upsilon_i|\upsilon_i > a)$.Then we have $E(\upsilon_i|\upsilon_i > a) = \mu_\upsilon + \sigma_\upsilon\frac{\phi(a)}{1-\Phi(\alpha)} = \mu_\upsilon + \sigma_\upsilon\lambda(a)$ and the variance is $V(\upsilon_i|\upsilon_i > a) = \sigma_\upsilon^2(1-\lambda(a)\left[\lambda(a)-a\right]) = \sigma_\upsilon^2\lambda(1-\delta(a))$. Hence,

$$\begin{aligned} E\left(w_i|I=1\right) &= x_i'\beta + \rho\sigma_\varepsilon\lambda_i(a), \\ Var\left(w_i|I=1\right) &= \sigma_\varepsilon 1 - \rho^2\lambda_i(a)\left(a - \lambda_i(a)\right). \end{aligned}$$

Since we do not know λ_i Heckman suggests a two step procedure to obtain the correct estimator. First, we use a probit model to fit a participation equation over the entire sample, and from this equation we recover the Mills ratio $\hat{\lambda}_i(a)$. Second we use this estimator as an explanatory variable when regressing wages on the characteristics of people with positive wages (which must be corrected for heteroskedasticity). The variance of the parameters estimators turns out to be consistent in the probit model since $\sqrt{n}\left(\hat{\lambda}_i - \lambda_i\right) \overset{A}{\backsim} N\left(0, \Sigma_i\right)$, with Σ_i the asymptotic variance-covariance matrix in the GLS estimation of the earning equation.

9.2 Maximum likelihood function

The two equations are

$$\begin{aligned} y_{1i} &= \alpha x_i + \varepsilon_{1i}, \\ y_{2i}^{*} &= \beta x_i + \varepsilon_{2i} \end{aligned}$$

with correlation $[\varepsilon_{1i}, \varepsilon_{2i}] = \rho_{12}$. The likelihood function in 12 can be decomposed in $L(\varepsilon_1)L(\varepsilon_{1i}/\varepsilon_2)$, concentrated over σ_2^2

$$\ln L = -\frac{N}{2}\ln(\frac{1}{N})\sum_i (y_{1i} - \alpha x_i)^2 + Tobit \ \log likelihood$$

define

$$\begin{aligned} d_0 &= 1 \text{ if } y_i \leq 0, \quad 0 \text{ otherwise } d_1 = 1 - d_0 \\ v_i &= y_{1i} - \alpha x_i \\ e_i &= y_{2i-}\beta x_i - \psi v_i \\ \psi &= \sigma_{12}/\sigma_1^2 \\ u_i &= \left(e_i - y_{i1}\right)/\omega \\ \omega &= \sigma_2^2\left(1 - \rho_{12}^2\right)^{1/2} \\ P &= \Phi\left(u_i\right) \end{aligned}$$

the Tobit term in the likelihood function is

$$-d_1 \left[\frac{1}{2} \ln \psi + \left(\frac{e_i}{\psi} \right)^2 \right] + d_0 \ln P$$